Atmospheric origins of extreme rainfall in the UK

By Andrew P. Barnes, Nick McCullen and Thomas R. Kjeldsen

Department of Architecture and Civil Engineering, University of Bath, Bath, UK

Abstract

Drawing causal relationships between physical processes and extreme rainfall is important for understand the extent of the impact these events can have. Current techniques fail to capture the causes of extreme rainfall, relying mostly on correlation techniques. This study extracts and classifies the storm trajectories of extreme rainfall events accross the UK using a novel trajectory dispersion model (HYSPLIT). These trajectories are then classified using three unsupervised classification techniques which were compared using a known cluster similarity measure, this resulted in the selection and application of the k-means method for identifying six of the key moisture pathways responsible for extreme rainfall. The most frequent pathway originates from the Atlantic and led to 60.58% of the extreme rainfall events, the remainder of the pathways tend to originate from the North Sea. Further to this, we identify the North Sea storms are the more likely to cause above average extreme events especially in Wales. A final comparison is made with the North-Atlantic Oscillation index where this study shows storms originating from the north and western Atlantic are more frequent during a positive NAO phase where as storms originating near the British Isles are more common in a negative phase.

1. Introduction

The impacts of extreme weather across the UK continue to have severe economic and social consequences. One key mechanism which can lead to a disaster is flooding often caused by extreme rainfall. A prime example of the costs associated with flooding can be seen in December, 2015 where record breaking levels of precipitation caused extensive flooding in Cumbria leading to at least 16,000 homes being flooded (MetOffice, 2016). Traditionally, design floods are calculated by fitting a statistical distribution to a sample of annual maximum events, assuming this distribution to be constant and that all observed events originate from the same underlying population. However, recent research has highlighted the importance of better understanding the underlying processes associated with individual events in order to: 1) build more robust models representing the existence of mixed populations (Kjeldsen *et al*, 2019), and 2) better understand how global climate change is likely to affect the type, magnitude and frequency of more localised distributions of extreme rainfall and floods (Bloschl *et al.*, 2018).

In this context it is important to develop new methods that will allow an objective classification of event types. For example, Lavers *et al.* (2011, 2012, 2013) highlights the role of large plumes of water rising from the tropics in the form of atmospheric rivers (ARs) which cause extreme levels of precipitation across the west coast of the UK.Recent literature has shown the application of atmospheric trajectory generation and classification as a potential tool for identifying the pathways of extreme events (Santos *et al.*, 2019; Tan *et al.*, 2017; Gimeno *et al.*, 2010) which would allow for the identification of

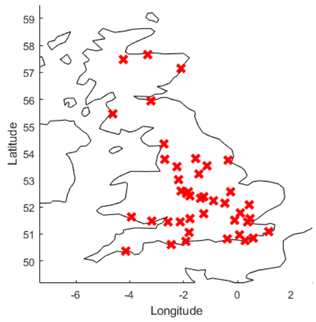

FIGURE 1. Cities selected for study.

atmospheric river type pathways to be identified as well as alternative pathways which cannot be explained by the AR theory. This study extracts and classifies the atmospheric trajectories for storms resulting in extreme rainfall across 42 cities in the UK. These classifications are then compared in terms of both magnitude, frequency, spatial dependency and their relation to the North Atlantic Oscillation.

2. Storm Track Extraction

Forty-two cities spread across the UK were selected as the initiation points for the identification of extreme rainfall events. The cities, as shown in Figure 1, were selected to cover all 10 hydroclimatic regions of Great Britain. For each city the daily annual maximum rainfall series (AMAX) was extracted from the CEH-GEAR dataset (Tanguy *et al.*, 2019). This data consists of gridded daily and monthly rainfall estimates between 1890-2017. Extracting the AMAX series from each of the 42 cities between 1960-2014 resulted in a total of 2598 extreme rainfall events to be analysed.

Next, the AMAX series were normalised between 0 and 1 to remove the dependency of each series on the local geographical and climatological features. The normalisation equation used is as given in Eq. (2.1), where $AMAX_c$ is the series of annual maximums for city c and $NAMAX_c$ is the series of normalised annual maximums for city c.

$$NAMAX_c = \frac{AMAX_c - min(AMAX_c)}{max(AMAX_c) - min(AMAX_c)}, c = 1, ..., 42 \tag{2.1}$$

For each of the 2598 AMAX events, the relevant atmospheric trajectories are generated from a trajectory dispersion model called HYSPLIT which requires an initial: longitude, latitude, altitude, extraction time, date and start time. The system also requires access to meteorological files covering the relevant dates. The extraction time indicates how many hours the model should trace the trajectory back. A 48 hour extraction time and NCEP/NCAR reanalysis data (Kalnay *et al.* 1996) were used in this study. To generate these trajectories HYSPLIT uses a combined Lagrangian and Eulerian approach which allow the relative calculation of the advection, diffusion and particle concentrations (Draxler and Hess, 1997; Draxler and Hess, 1998; Draxler, 1999; Stohl and James, 2004).

To extract the trajectories for each of the 2598 extreme events the HYSPLIT model was initialised with two altitude starting points (500 and 1000m above sea level) and two start times per event-day (09:00 and 18:00). This resulted in 9488 trajectories. All trajectories were normalised to their initiation points, thereby removing the dependency of the trajectory on the initiation point as shown in Eq. 2.2.

$$NT_n = T_n - T_0 \tag{2.2}$$

Each trajectory (T) consists of 49, $2 * 1$ vectors representing the latitude and longitude of T at each timestep n. These points are normalised according to the trajectories initiation point T_0 to produce a new 49-point normalised trajectory (NT).

3. Pathway Classification

Three unsupervised classification methods are compared for their suitability. The three methods chosen were K-means, Self-organising maps (SOMs) and a linkage method using the ward approach. For each of these methods a predetermined number of intended output classifications is required, to optimise each method a classification was carried out with a number of output classes ranging from 2 to 50. It was considered that a number of classifications beyond this point would become unwieldy for a qualitative analysis.

Next, to allow a comparison of the suitability of each method the Davies-Bouldin (Davies & Bouldin, 1979) index is used to both optimise the number of clusters for each method and to compare the resulting optimum models. This index was used as it provides an accurate indicator of cluster similarity and distance (Halkidi *et al.*, 2001). An alternative to this is the Calinski-Harabasz (Caliński & Harabasz, 1974) method. However this method performs better when clusters are well separated, an assumption which cannot be imposed on the trajectory data. The Davies-Bouldin index measures the average similarity between each cluster C_i where $i = 1, ..., N$ and its most similar cluster C_j. The similarity measure R_{ij} as defined in Eq. 2.3, where S_i and S_j correspond to the cluster diameter (average distance of points to the cluster center), and M_{ij} is the distance between the centroid of cluster i and j.

$$R_{ij} = \frac{S_i + S_j}{M_{ij}} \tag{3.1}$$

Using this similarity measure the Davies-Bouldin index is then defined as below where N is the number of clusters present. As DB_{index} increases the similarity between clusters is increasing, hence a lowerDB_{index} indicates better cluster separation.

$$DB_{index} = \frac{1}{N} \sum_{i=1}^{N} max(R_{ij} | i \neq j) \tag{3.2}$$

Figure 2 shows the resulting Davies-Bouldin indices and optimal number of clusters (optimal value of N) for each of the three methods selected. Each optimal N values identified are in the lower end of the possible values of N with $N = 4$ for SOMs, $N = 5$ for the linkage method and $N = 6$ for k-means. Consequently, it was decided to take a single classifier from each of the three methods using their optimal number of output classifiations (clusters) namely SOM_4, LNK_5 and KME_6.

4. Results

A random selection of extracted trajectories, including the centroid, is shown in Figure 3 for each classification as defined by the three optimal models. In addition Table 1 gives the proportion of the trajectories classified under each class. The proportions allocated to each classification appear to get lower when moving from a model with only four

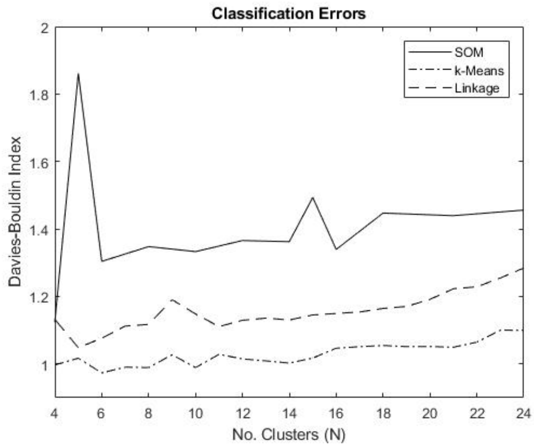

FIGURE 2. Davies-Bouldin Index for each of the three selected clustering methods with a varying number of initiation clusters.

Classifier	Classification Proportions (%)					
	1	2	3	4	5	6
SOM_4	36.19	10.58	7.53	45.70		
LNK_5	8.36	27.36	19.65	10.76	33.87	
KME_6	13.92	25.63	9.19	12.60	24.71	13.95

TABLE 1. Proportion of trajectories classified under each classification per model.

output clusters (SOM_4) to one with six (KME_6). However, this is to be expected as more clusters allow further separation of similar trajectories.

The storms originating from the southern Atlantic, such as clusters 4 and 5 in the k-means model ($KME_6^{4,5}$), cluster 1 in the linkage model (LNK_5^1) and a small proportion of cluster 1 in the SOM model (SOM_4^1) tend to follow a similar trajectory to those of atmospheric rivers as identified by Lavers *et al.* (2011, 2012), and hence could potentially be attributed to the occurrence of such phenomena. Taking the proportions of $KME_6^{4,5}$ it would indicate that events caused by these atmospheric rivers could make up 37.31% of the extreme events. Further to this, the second most frequent sub-set of classes originate from the North Sea where classes such as SOM_4^4 ,LNK_5^5 and $KME_6^{1,2}$, 2contain 36.19%, 33.87% and 39.55% of all events, respectively. This further illustrates the reliance of extreme events on both the Atlantic and North Sea atmospheric processes. Comparing the length of the cluster centroids shows the Atlantic based storms namely $SOM_4^{1,2,3}$, $LNK_5^{1,2,3,4}$ and $KME_6^{3,4,5,6}$ travel a much longer distance than their North sea counterparts (SOM_4^4 ,LNK_5^5 and $KME_6^{1,2}$). This suggests a stronger driving force moving the air parcels above the Atlantic Ocean at a much higher speed than those travelling from the North Sea.

The spatial distribution of the classifications was considered using only KME_6 as these pathways are comparable to those in the other models, and this classifier had the lowest DB index as discussed earlier. Here a contrast between the eastern and western regions can be observed with regards to KME_6^2 pathways, where the pathway leading from the North Sea westwards affect the eastern regions more than those in the west. For comparison, 42% of the trajectories from the North East of England belong to KME_6^2

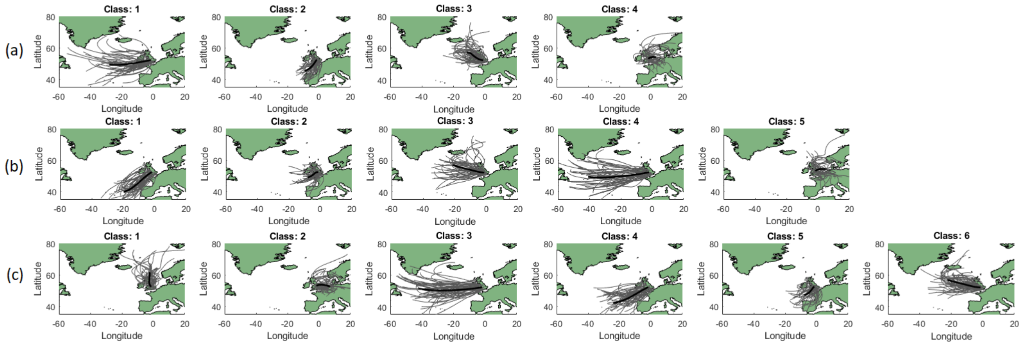

FIGURE 3. Examples of the trajectories classified under each class within the three classifiers used: SOM_4 (a), LNK_5 (b) and KME_6 (c).

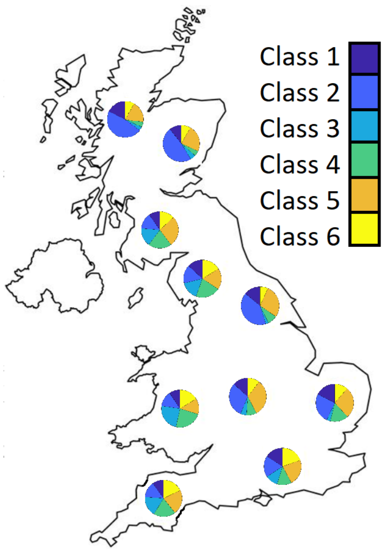

FIGURE 4. Proportion of each KME_6 storm classification per hydroclimatic region.

where as only 16% of the trajectories leading to events in the North West belong to KME_6^2. One cause of this pattern could be the existence of the Pennines and the Peak district which prevent the storms from the east carrying moisture to the western side of Great Britain. Furthermore, the western regions have a larger portion of KME64 type trajectories; specifically North West England, South West England and Wales have 21%, 19% and 24% of their samples classified as KME_6^4. The East of England also shows a large proportion of KME_6^4 trajectories which total 15% in the region. This further highlights the importance of the geography of Great Britain, as these storms are unable to penetrate the elevated terrain of the Pennines and the peak district as often as they are able to make it across the mostly flat regions in the south. Finally, the most frequent pathway resulting in extreme events in Wales appears to come from the South Atlantic (KME_6^4). A similar trend is seen in South West England which corresponds to findings by Svensson and Jones (2004) who identified north-easterly moving storm tracks to have an association with high-river flows in these areas.

Next, the proportion of above average extremes was investigated for each of the classes in KME_6. Table 2 shows for each class the proportion of events with a magnitude greater

Classifier	Classifications					
	1	2	3	4	5	6
SOM_4	29.9	31.9	38.8	46.8		
LNK_5	30.0	42.3	33.2	30.1	46.2	
KME_6	47.1	47.6	30.1	38.3	41.1	30.9

Table 2. Proportion of events in each classification with a magnitude ¿ 1.

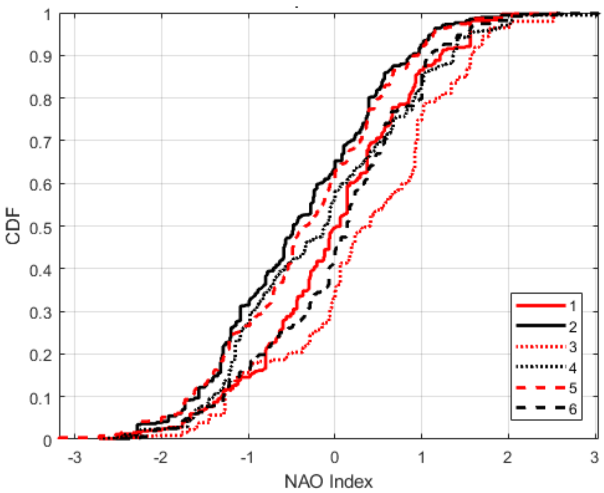

Figure 5. Empirical cumulative distribution function for the frequency of each classification for a given NAO index.

than the normalised mean of 1 (normalised according to Eq. 2.1). The classes identified as originating from the North Sea SOM_4^4, LNK_5^5 and $KME_6^{1,2}$) all contain the highest proportion of above-mean events where as those from the deep-west Atlantic (SOM_4^4, LNK_5^4 and KME_6^3) have the lowest proportions. Regional difference in these magnitude variations are also matched within each region with a few key differences, 61.3% of KME_6^2 trajectories cause above average extremes in Wales where as these types only cause 38% of above average extremes in South West England.

5. Linking trajectory classifications to North-Atlantic Oscillation

By performing a linear regression between the magnitude of the trajectories in each classification and their relevant NAO indices we have found no significant correlation between these two variables. However, Figure 5 shows the cumulative proportion of events for each classification given the NAO indices of each.

These results show KME_6^3 has the highest proportion of events falling within a positive NAO phase totalling 67% followed by KME_6^6 at 59%. This indicates events caused by pathways leading from the deep-West and North Atlantic become more frequent during a positive NAO phase. Furthermore, during the negative NAO phase it is the KME_6^2 and KME_6^5 pathways which are the most prominent with 63% and 60% of their respective trajectories falling during this negative phase. On visual inspection it appears pathways in KME_6^2, 5originate from close to the British Isles and form circular patterns which could indicate the existence of cyclonic activity in the south. Finally, inspecting the empirical CDF pattern of pathway KME_6^3 it appears to spike in frequency during an NAO index of 0 and 1 which could also indicate the triggering of an atmospheric event at these levels.

6. Conclusions

This study has developed a new framework for identifying the key-moisture pathways leading to extreme rainfall events in the United Kingdom through the classification of the trajectories of the storms prior to the occurrence of extreme rainfall events. This is accomplished through the optimisation of three unsupervised classification methods using the Davies-Bouldin index to select an optimal number of output nodes and analysing the resulting trajectory classes. The key findings were as follows:

(*a*) Atlantic and North Sea storms form two subsets of the resulting classifications from all three models with 60.58% of the trajectories in the k-Means classifier (KME_6) belonging to the Atlantic pathways and 39.42% originating from the North Sea.

(*b*) Classes in each of the three models match the pathways of atmospheric rivers, a known cause of high levels of rainfall in the UK (Lavers*et al.*, 2011, 2012).

(*c*) The North Sea storms cause a higher proportion of events across the easter regions of Great Britain where as the South Atlantic storms are more commonly associated with those across the West.

(*d*) Westerly storms from the North Sea is the most likely pathway to cause above average extreme events in Wales in comparison with the other surrounding regions.

(*e*) A comparison of the magnitude of events in each class and the North Atlantic Oscillation showed no significant relationship between these two variables.

(*f*) The frequency of storm originating from the North and West Atlantic increases with a positive NAO index whereas short storms originating near the British Isles become more frequent with a negative NAO phase.

These results highlight the importance atmospheric pathways have on the temporal and spatial distribution of extreme rainfall events across the UK. The resulting classifications presented spatial and magnitude difference across all ten hydroclimatic regions but have also opened new questions regarding the atmospheric processes which lead to these events via the pathways identified.

REFERENCES

Blosch, G., Hll, J., Parajka, J. ... 2017. Changing climate shifts timing of European floods. Science, 357, 6351, 588-590.

Cabinet Office., DEFRA., Gummer, B., Leadsom, A. 2016. National Flood Resilience Review. [online] Available at: https://www.gov.uk/government/publications/national-flood-resilience-review [Last accessed: 04/07/2019]

Cox, D., Hunt, J., Mason, P., Wheater, H., Wolf, P., Robson, A.J. 2002. Evidence for trends in uk flooding. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 360, 1327-1343.

DEFRA. 2012. UK Climate Change Risk Assessment. International Journal of Climate Change Strategies and Management, 4, 2.

Draxler, R., 1999. HYSPLIT4 user's guide. NOAA Tech. Memo. ERL ARL-230, (June), NOAA Air Resources Laboratory, Silver Spring, MD.

Draxler, R.R. and Hess, G.D., 1997. Description of the HYSPLIT4 modeling system. Techreport.

Draxler, R.R. and Hess, G.D., 1998. An Overview of the HYSPLIT4 Modelling System for Trajectories, Dispersion, and Deposition. Australian Meteorological Magazine, 47, 295–308.

Government Office for Science. 2004. Future Flooding. [online] Available at: https://www.gov.uk/government/publications/future-flooding [Last accessed: 04/07/2019]

Kjeldsen, T.R., Ahn, H., Prosdocimi, I. and Heo, J.H., 2018. Mixture Gumbel models for extreme series including infrequent phenomena. Hydrological Sciences Journal, 63(13-14), pp.1927-1940.

Lavers, D.A., Allan, R.P., Wood, E.F., Villarini, G., Brayshaw, D.J., Wade, A.J. 2011. Winter Floods in Britain are Connected to Atmospheric Rivers. Geophysical Research Letters, 38, 23.

Lavers, D.A., Villarini, G., Allan, R.P., Wood, E.F., Wade, A.J. 2012. The Detection of Atmospheric Rivers in Atmospheric Reanalyses and their Links to British Winter Floods and the Large-scale Climatic Circulation. Journal of Geophysical Research: Atmospheres, 117, D20.

MetOffice. 2016. Flooding in Cumbria December 2015. [online] Available at: https://www.metoffice.gov.uk/climate/uk/interesting/december2015. [Last accessed: 04/07/2019]

Salas, J. D., Obeysekera, J., and Vogel, R.D., 2018. Techniques for Assessing Water Infrastructure for Nonstationary Extreme Events: A Review. Hydrological Sciences Journal 63, 3, 325–352.

Stohl, A. and James, P., 2004. A Lagrangian Analysis of the Atmospheric Branch of the Global Water Cycle. Part I: Method Description, Validation, and Demonstration for the August 2002 Flooding in Central Europe. Journal of Hydrometeorology, 5 (4), 656–678.

Svensson, C., Jones, D.A. 2004. Dependence between sea surge, river flow and precipitation in south and west Britain. Hydrology and Earth System Sciences, 8 (5), 973-992.

Tanguy, M.; Dixon, H.; Prosdocimi, I.; Morris, D.G.; Keller, V.D.J. (2019). Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2017) [CEH-GEAR]. NERC Environmental Information Data Centre. https://doi.org/10.5285/ee9ab43d-a4fe-4e73-afd5-cd4fc4c82556